

Creating rich shareable metadata: The DLF Aquifer MODS implementation guidelines

Sarah L. Shreeves
University of Illinois at Urbana-Champaign

Using Metadata Standards in Digital Libraries:
implementing METS, MODS,
PREMIS and MIX – Sponsored by the LITA
Standards Interest Group

ALA Annual – Washington, D.C.
June 25 2007



Outline

- Context – DLF and Aquifer
- Guidelines
- What is Shareable metadata?
- Guidelines
- What are we working on now?

What is the Digital Library Federation?

... an international association of libraries and allied institutions.

Mission:

to enable new research and scholarship of its members, students, scholars, lifelong learners, and the general public by developing an international network of digital libraries.

DLF Aquifer Initiative

<http://www.diglib.org/aquifer/>

Purpose is to promote effective use of distributed digital library content for teaching, learning, and research in the area of American culture and life.

Bringing distributed content together and pushing it out to scholars and researchers

Awarded \$816,000 from the Mellon Foundation for *DLF Aquifer Development for Interoperability Across Scholarly Repositories: American Social History Online*

Initial aggregation at: <http://quod.lib.umich.edu/a/aquifer/>

June 25, 2007

Currently 230,340 MODS records from 5 institutions

Indiana University Digital Library
Program: Sam DeVincent Collection of
American Sheet Music
6 records

Library of Congress Digitized Historical
Collections: Bain Collection
(photographs)
5 records

Library of Congress Digitized Historical
Collections: Brady-Handy Collection
(Photographs)
2 records

Library of Congress Digitized Historical
Collections: Cartoon Prints, American,
1766-1876
1 record

Library of Congress Digitized Historical
Collections: Civil War Photographs
1 record

Library of Congress Digitized Historical
Collections: LC Early Motion Pictures
2 records

Library of Congress Digitized Historical
Collections: LC Maps
5 records

Library of Congress Digitized Historical
Collections: LC Panoramic Photographs
3 records

Library of Congress Digitized Historical
Collections: LC Sheet Music
20 records

Library of Congress Digitized Historical
Collections: Sheet Music

Record 1 of 52

[add to bookbag](#)

Title	Annual Banquet, National Hardwood Lumber Dealers Assn., Memphis, Tenn
Author/Creator	Geo. R. Lawrence Co, copyright claimant.
Type	still image
Issue Date	c1906; 1906
Genre	graphic; Group portraits.; Portrait photographs.; Banquet camera photographs.; Gelatin silver prints.
Physical Description	graphic; 1 photographic print : gelatin silver ; 18 x 28.5 in.
Subject: Topical	Banquets; People
Subject: Hierarchical	United States; Tennessee; Memphis
Subject: Proper Names	National Hardwood Lumber Association
Note	Copyright claimant's address: Washington, DC.
Note	No known restrictions on publication.
Note	Copyright deposit; Geo. R. Lawrence Co.; May 11, 1906.
Identifier	LC-USZ62-41405 DLC; hdl:loc.pnp/pan.6a34842
Classification	LOT 5787 no. 23
Location	Library of Congress Prints and Photographs Division Washington, D.C. 20540 USA
URL	http://hdl.loc.gov/loc.pnp/pan.6a34842
Rights	No known restrictions on publication.
Data Contributor	Library of Congress Digitized Historical Collections: LC Panoramic Photographs



Record 2 of 52

[add to bookbag](#)

Title	Back to my sweetheart
Alternative Title	Softly o'er the mountain fall the dear old golden moon

Background

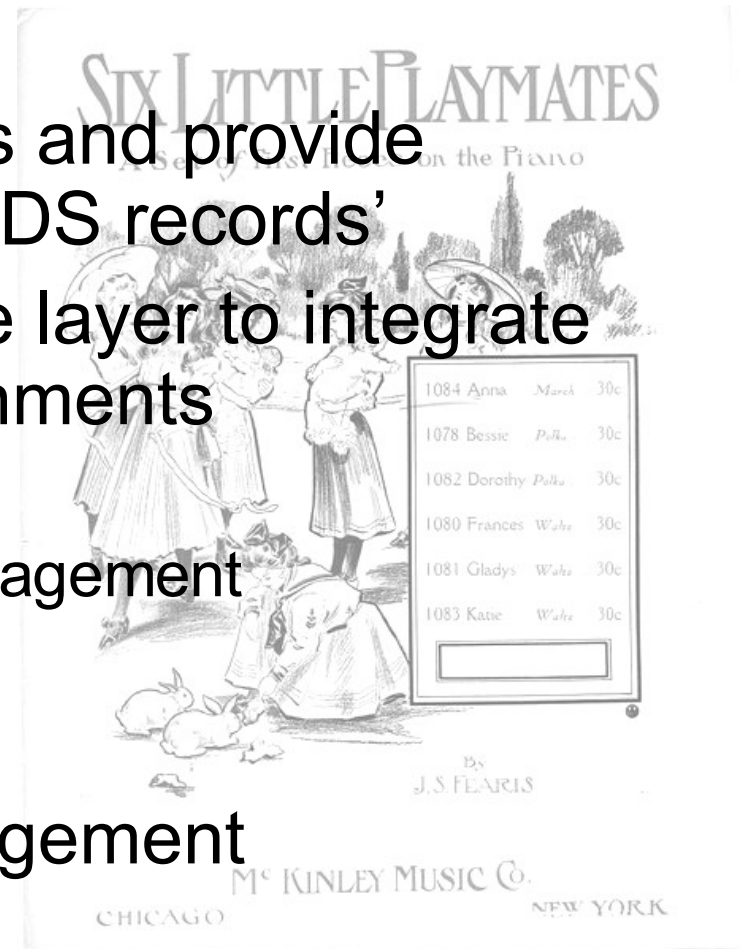
- OAI PMH allows metadata to be harvested from multiple data providers
 - Usually metadata is brought together into aggregations (like OAIster)
- OAI PMH requires use of simple Dublin Core but allows use of other metadata schemas
 - Most data providers only provide Dublin Core; most aggregators only harvest Dublin Core

Difficulties in Building Services in Aggregations

- Use of Dublin Core often problematic
 - Lack of semantic complexity
- Metadata authors often are not thinking about sharing metadata outside of their local contexts
- Metadata between different collections inconsistent
- Few user studies of aggregations
- Building browse, targeted search, facets, etc... has been very, very, very difficult for aggregators

What's different about Aquifer's American Social History Online?

- Only harvest MODS records and provide guidance for 'shareable MODS records'
- Interoperability / middleware layer to integrate collections into local environments
 - Course management system
 - Portal with link to citation management
 - Commercial search service
 - Federated search
- User studies/end user engagement



DLF Aquifer Implementation Guidelines for Shareable MODS Records

Developed by Aquifer Metadata WG

- Sarah L. Shreeves (University of Illinois at Urbana-Champaign): 2005 - (Chair 2005 – 2006)
- Jenn Riley (Indiana University): 2005 – (Chair 2007 -)
- Laura Akerman (Emory University): 2006 -
- John Chapman (University of Minnesota): 2005 -
- Melanie Feltner-Reichert (University of Tennessee): 2006 -
- Bill Landis (California Digital Library): 2005 -
- Tracy Meehleib (Library of Congress): 2006 -
- Elizabeth Milewicz (Emory University): 2005 - 2006
- David Reynolds (Johns Hopkins University): 2005 -
- Gary Shawver (New York University): 2005 –

DLF Aquifer Implementation Guidelines for Shareable MODS Records

- To help create rich, shareable metadata records that is consistent and coherent
- Meant for shared records, not local implementations
- Focused on humanities and cultural heritage materials
- Based on the *DLF / NSDL Best Practices for Shareable Metadata*

http://www.diglib.org/aquifer/dlfmodsimpmentationguidelines_finalnov2006.pdf

What is shareable metadata?

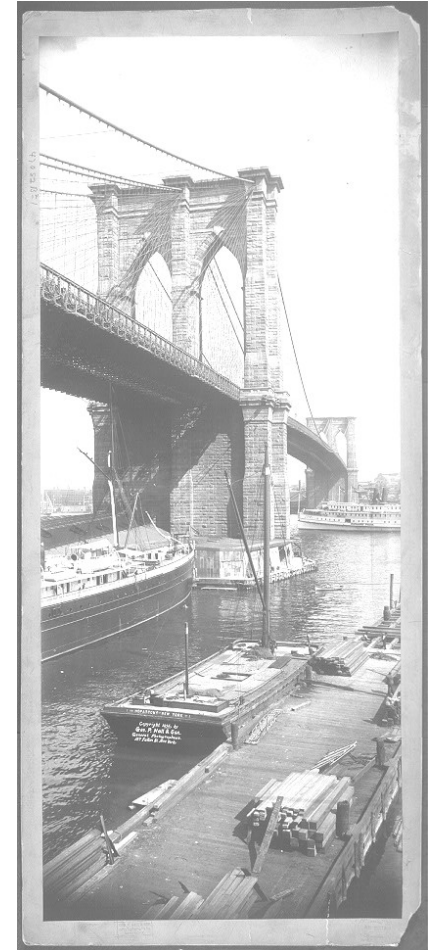
- Is **quality** metadata (see Bruce and Hillmann)
- Promotes **search interoperability**
“the ability to perform a search over diverse sets of metadata records and obtain meaningful results.”
(Priscilla Caplan)
- Is **human understandable** outside of its local context
- Is **useful** outside of its local context
- Is **machine processable**

Metadata as a view of a resource

- No monolithic one-size-fits-all metadata record
- The view might be different depending on *use* and *audience* as well as format, content, and context
 - Content standard is a view
 - Metadata standard is a view
 - Vocabulary used is a view

Format of the Guidelines

- Summary of Requirements and Recommendations
- Definition from LC's MODS User Guidelines
- Discussion of Use
- Examples
- Use by Aggregators
- Mapping to Dublin Core
- Relationship to DLF/NSDL Best Practices



Required elements

<titleInfo><title>

<typeOfResource>

<originInfo> and at least one date subelement with
keydate="yes"

<language> if language primary to resource

<physicalDescription><digitalOrigin>

<physicalDescription><internetMediaType>

<subject> if applicable

<location><url usage="primary display">

**<accessCondition type="use and
reproduction">**

<recordInfo><languageOfCataloging>

<recordInfo><languageTerm>

Recommended elements

<name><namePart>

<genre>

<abstract>

<tableOfContents> if applicable

<targetAudience> if applicable

<note> if applicable

<relatedItem> if applicable

<identifier>

<part> if applicable

Optional elements

<classification>

<extension>

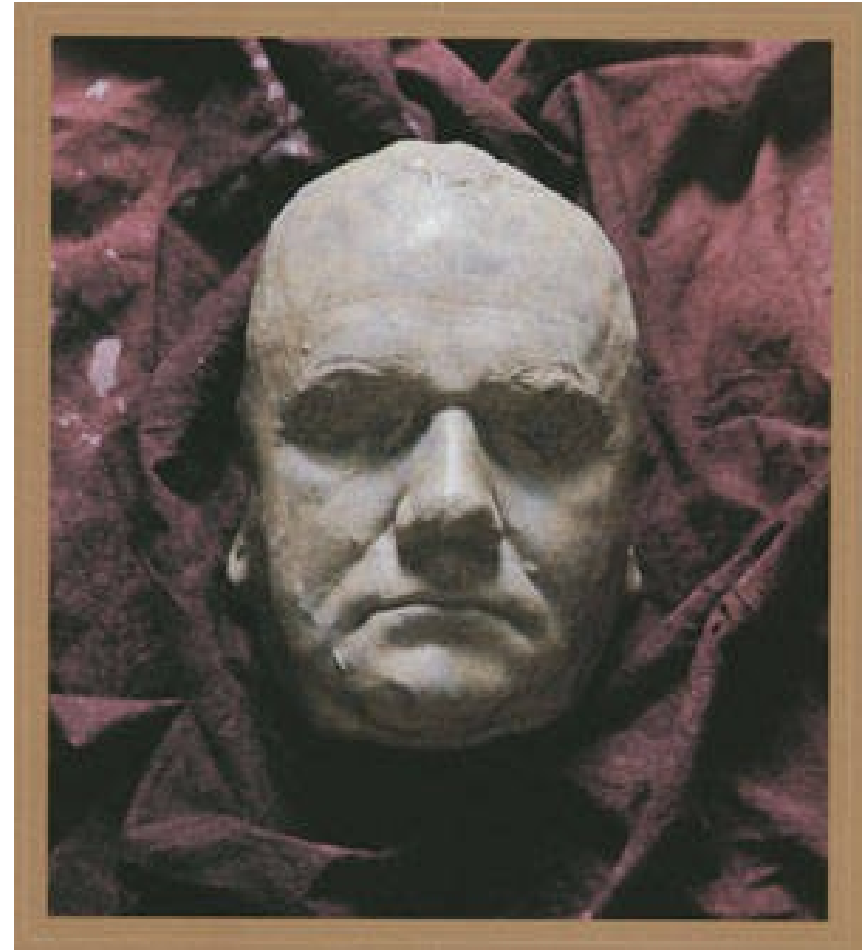


Levels of Adoption - Tied to user tasks

- 1. Minimum for participation: Allows users to cite the resource**
- 2. Minimum for doing anything useful: Allows users to perform basic searches and filtering**
- 3. Allows more advanced functionality: Allows users to browse and group search results**
- 4. Adopt all required guidelines (and some recommended): Allows users to perform more precise searches**
- 5. Completely adopt all recommendations: Allows users to effectively evaluate resources**

Example – Level 3

Title: Life Mask of Stephen A. Douglas
Name (authority=naf): Leonard Wells
Role: Sculptor
Type of Resource: three dimensional object
Genre: life masks
Genre: casts (sculpture)
Date created (w3cdtf keyDate="yes"): 1858
Form authority="gmd": art original
InternetMediaType: image/jpeg
Subject: Douglas, Stephen Arnold 1813-1861
URL usage="primary display"
access="object in context":<http://images.library.uiuc.edu:8081/u?/tdc,252>
AccessCondition
type="useAndReproduction": Use of this public-domain resource is unrestricted.



What else are we working on?

- Release of the level of adoptions
- MARC → MODS stylesheet
 - Currently in testing
- Validation tool for MODS records against the full guidelines
 - Still in development
- Also, Aquifer is interested in collections (from non-DLF institutions as well as DLF institutions) – contact Katherine Kott, Director of the Aquifer Initiative.



Contact Information

Sarah Shreeves

sshreeve@uiuc.edu

217.244.3877

References:

Bruce, T.R. and D.I. Hillmann. (2004) "The continuum of metadata quality: defining, expressing, exploiting," in *Metadata in Practice*, Ed. by Diane Hillmann and Elaine Westbrooks. Chicago: ALA Editions.

Tennant, R. (2004) "Metadata's bitter harvest," *Library Journal* 129, no.12. <http://www.libraryjournal.com/article/CA434443.html>

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 License